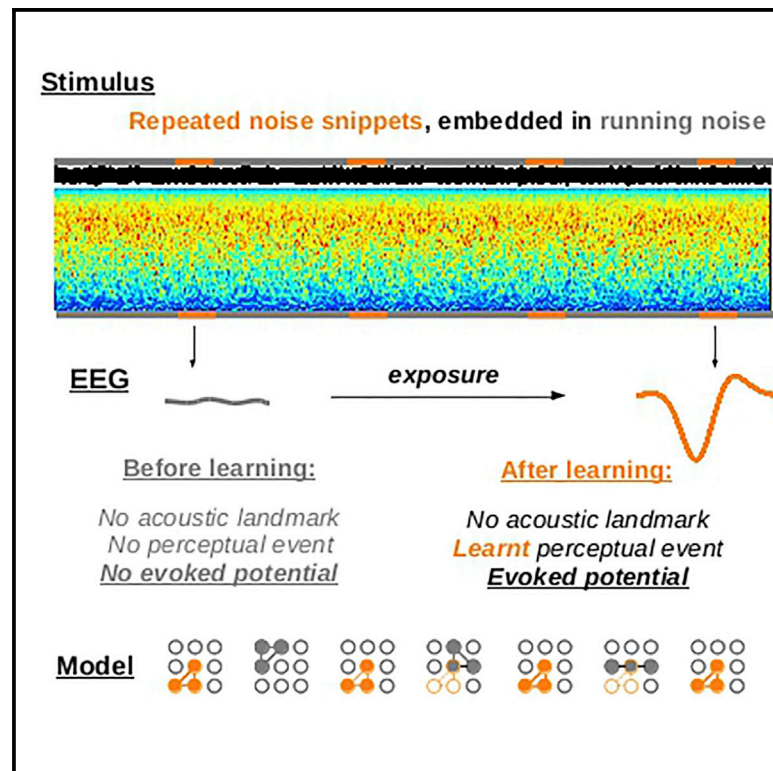# Current Biology

Perceptual Learning of Acoustic Noise Generates Memory-Evoked Potentials

Graphical Abstract

Authors
Thomas Andrillon, Sid Kouider, Trevor Agus, Daniel Pressnitzer

Correspondence
thomas.andrillon@ens.fr (T.A.),
daniel.pressnitzer@ens.fr (D.P.)

In Brief
Andrillon et al. investigated how human listeners manage to learn complicated random auditory noises after only a few exposures. They showed that learning was tracked in real time by the emergence of novel auditory potentials. These neural responses could signal the extremely rapid formation of sharp selectivity to subtle acoustic patterns.

Highlights

- The auditory learning of noise induced fast qualitative changes in EEG signals

- Behavioral memory traces were mirrored by time-locked sensory potentials

- Rapid plasticity seems able to create sharp selectivity to complex auditory cues

CellPress

Perceptual Learning of Acoustic Noise Generates Memory-Evoked Potentials

Thomas Andrillon,[1,2,3,*] Sid Kouider,[1,2,4] Trevor Agus,[2,5,6] and Daniel Pressnitzer[2,5,*]

[1]Laboratoire de Sciences Cognitives et Psycholinguistique (UMR8554, ENS, EHESS, CNRS), 29 rue d'Ulm, 75005 Paris, France
[2]Département d'Études Cognitives, École Normale Supérieure – PSL Research University, 29 rue d'Ulm, 75005 Paris, France
[3]École Doctorale Cerveau Cognition Comportement, Université Pierre et Marie Curie, 1 Place de Jussieu, 75005 Paris, France
[4]Science Division, Department of Psychology, New York University Abu Dhabi, Saadiyat Island, PO Box 129188, Abu Dhabi, UAE
[5]Laboratoire des Systèmes Perceptifs, CNRS UMR 8248, 29 rue d'Ulm, 75005 Paris, France
[6]Sonic Arts Research Centre, School of Creative Arts, Queen's University Belfast, Cloreen Park, BT9 5HN Belfast, UK
*Correspondence: thomas.andrillon@ens.fr (T.A.), daniel.pressnitzer@ens.fr (D.P.)
http://dx.doi.org/10.1016/j.cub.2015.09.027

SUMMARY

Experience continuously imprints on the brain at all stages of life. The traces it leaves behind can produce perceptual learning [1], which drives adaptive behavior to previously encountered stimuli. Recently, it has been shown that even random noise, a type of sound devoid of acoustic structure, can trigger fast and robust perceptual learning after repeated exposure [2]. Here, by combining psychophysics, electroencephalography (EEG), and modeling, we show that the perceptual learning of noise is associated with evoked potentials, without any salient physical discontinuity or obvious acoustic landmark in the sound. Rather, the potentials appeared whenever a memory trace was observed behaviorally. Such memory-evoked potentials were characterized by early latencies and auditory topographies, consistent with a sensory origin. Furthermore, they were generated even on conditions of diverted attention. The EEG waveforms could be modeled as standard evoked responses to auditory events (N1-P2) [3], triggered by idiosyncratic perceptual features acquired through learning. Thus, we argue that the learning of noise is accompanied by the rapid formation of sharp neural selectivity to arbitrary and complex acoustic patterns, within sensory regions. Such a mechanism bridges the gap between the short-term and longer-term plasticity observed in the learning of noise [2, 4–6]. It could also be key to the processing of natural sounds within auditory cortices [7], suggesting that the neural code for sound source identification will be shaped by experience as well as by acoustics.

RESULTS

We used an experimental paradigm where listeners learned exemplars of acoustic noise [2, 5, 6, 8]. Although noise is not representative of natural sounds, it is a unique tool to probe auditory

[2] or even visual [9, 10] perceptual learning. First, noise lacks semantic content, thus revealing pure perceptual learning. Also, there is normally no prior exposure to a specific noise exemplar. Finally, noise exemplars can contain tens of thousands of random samples, with no sample-to-sample predictability, pushing any learning mechanism to the extreme.

The surprising ability of listeners to learn meaningless, random patterns is also relevant to the long-standing debate about the nature of experience-dependent changes in the brain: are such changes distributed or local [1, 11]? A distributed code posits subtle changes in a whole population of neurons, with the functional benefits appearing only at the population level. For audition, perceptual learning has indeed been observed through distributed changes in "tonotopic" frequency maps, using pure tones [12, 13]. As noise has a flat spectrum on average, a distributed code could not rely on tonotopy, but it could possibly recruit more-complex timbre maps [14, 15]. In contrast, a local code posits dramatic changes at the single-neuron or small-network level, with small neural populations expressing the full benefit of learning [7, 16]. For noise, such a localized code could apply, if repeated exposure to a random pattern created a form of ultra-selective response [17, 18].

Recently, Luo et al. [5] applied magnetoencephalography (MEG) to the noise-learning paradigm. They found that noise learning induced stable phase patterns in brain neural responses, as measured by inter-trial phase coherence (ITPC) in the 3–8 Hz theta range. Those results were interpreted as a distributed and holistic learning process [19, 20]. Strikingly, there was no effect of perceptual learning on event-related potentials (ERPs), which further advocated against a local process: if listeners learned isolated perceptual events within the noise, those should be accompanied by ERPs [21]. However, the local account may be reprieved with an additional hypothesis. If learned features were local but idiosyncratic, and thus activated at random times across listeners for the same noise [2, 22, 23], then the associated ERPs would be impossible to observe on average, whereas ITPC would remain high. We devised a variant of the noise-learning paradigm to test this hypothesis.

Behavioral Measures of Perceptual Learning: Diffuse and Compact Conditions

For half of the experiment (Figure 1A, left), a standard noise-learning paradigm was used [2, 5]. Participants were
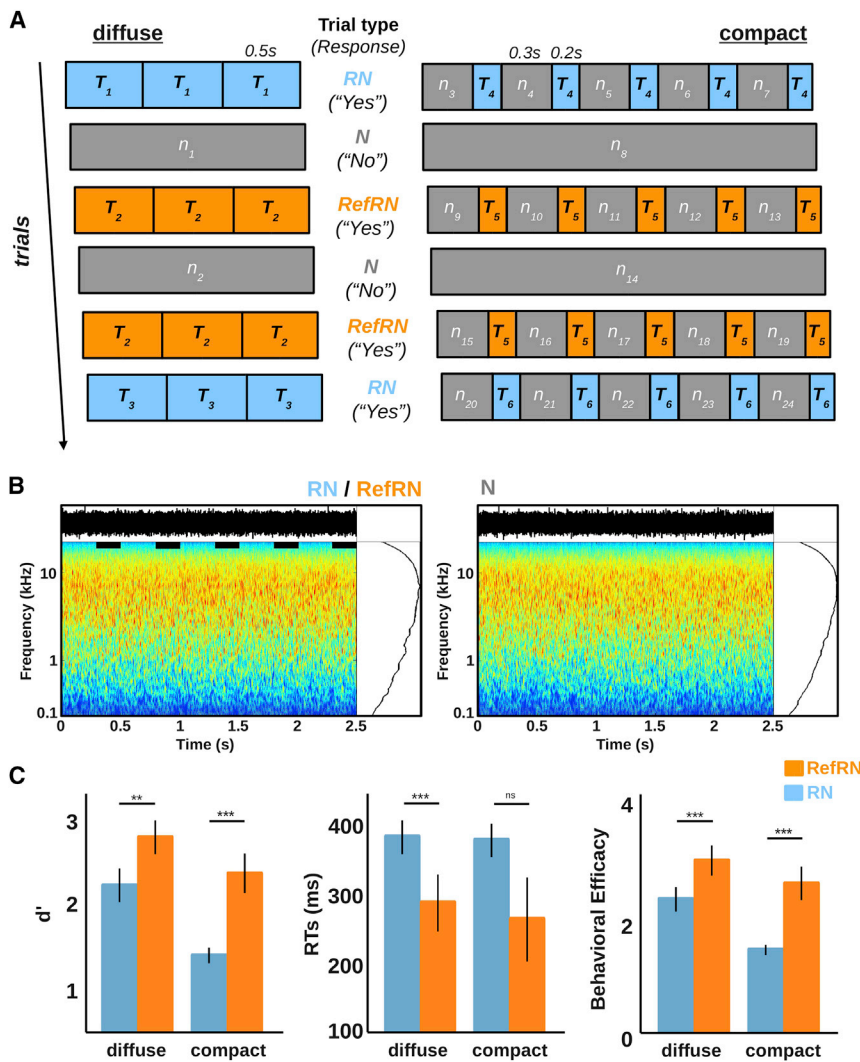
Figure 1. Experimental Procedure and Behavioral Results

(A) Stimuli and experimental design. In the "diffuse" condition (left), participants had to discriminate between no-repeat trials, made of 1.5 s of running white-noise (N), and with-repeats trials, made by concatenating, without any discontinuity, three identical 0.5-s-long snippets of noise. For repeated noise trials (RN), a new snippet of noise was drawn for each trial. For reference repeated noise trials (RefRN), the exact same noise snippet re-occurred not only within a trial but also across trials (illustrated here with target snippet T_2). In the "compact" conditions, the task was the same, but repeated snippets were shorter (0.2 s) and concatenated, without any discontinuity, with 0.3-s running white-noise fillers. Compact trials lasted 2.5 s and included five 0.5-s-long partially repeating epochs (Supplemental Experimental Procedures).

(B) Typical output of a peripheral auditory model (spectro-temporal excitation pattern [24]; see Supplemental Experimental Procedures for details) for repeated (RN/RefRN) and unrepeated (N) stimuli. The compact condition was used for the simulations. There are no obvious landmarks for the repeated stimuli in the waveform (top), average excitation pattern (right), nor spectro-temporal excitation pattern (main panel).

(C) Behavioral performance, averaged over n = 42 blocks. Three measures were computed: sensitivity d'; reaction times (RTs); and behavioral efficacy combining d' and RTs (Supplemental Experimental Procedures). Error bars denote SEM across blocks. Stars indicate the significance level for the RefRN versus RN comparisons (paired u test; ns: $p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.005$). RTs were faster for RefRN in the compact, but not diffuse, condition (paired u test: $p = 0.21$). A better performance (higher d' and behavioral efficacy; faster RTs) for RefRN compared to RN summarizes the amount of perceptual learning.

instructed to discriminate trials containing continuous white noise (N) from trials made of the seamless concatenation of three copies of a 0.5-s-long white-noise snippet (repeated noise [RN]). Without participants' knowledge, a third type of trial was introduced: one particular instance of RN (reference RN [RefRN]) re-occurred, identically, over 16 trials throughout a block. A higher repetition-detection performance for RefRN relative to RN indicates perceptual learning [2]. For the other half of the experiment, a change was introduced in the structure of trials containing repetitions: a shorter 0.2-s-long white-noise snippet was repeated but seamlessly concatenated, between repetitions, to 0.3-s-long fresh noises (Figure 1A, right). Thus, the temporal window over which repetition detection and perceptual learning could occur was restricted to 0.2 s. This would induce less temporal variability in putative local EEG markers. We use the terms "diffuse" for fully repeating noise [2, 5] and "compact" for partially repeating noise [25, 26].

Behavioral measures showed clear evidence of perceptual learning in both conditions (Figure 1C). First, signal-detection

analysis [27] showed a better d' sensitivity for RefRN compared to RN. This difference between RefRN and RN was absent at the beginning of blocks, that is, before learning of RefRN could occur (Figure S1A). Reaction times (RTs) were faster for RefRN than RN in the diffuse condition, and a higher accuracy was associated with faster responses for RefRN in both diffuse and compact conditions (Figure S1B). This suggests that both d' and RTs indexed perceptual learning. We combined d' and RTs in a "behavioral efficacy" index (BE) (Supplemental Experimental Procedures). The compact condition led to lower BEs (Friedman test; $p < 0.001$), showing that this condition was more difficult overall. However, the amount of learning, as measured by the BE difference between RefRN and RN, was the same across conditions (paired u test; $p = 0.46$).

Electrophysiological Markers of Learning

EEG was recorded while participants performed the task, and analyses were restricted to the sensors most responsive to auditory stimuli (Supplemental Experimental Procedures). Following [5], we investigated three possible neural markers of
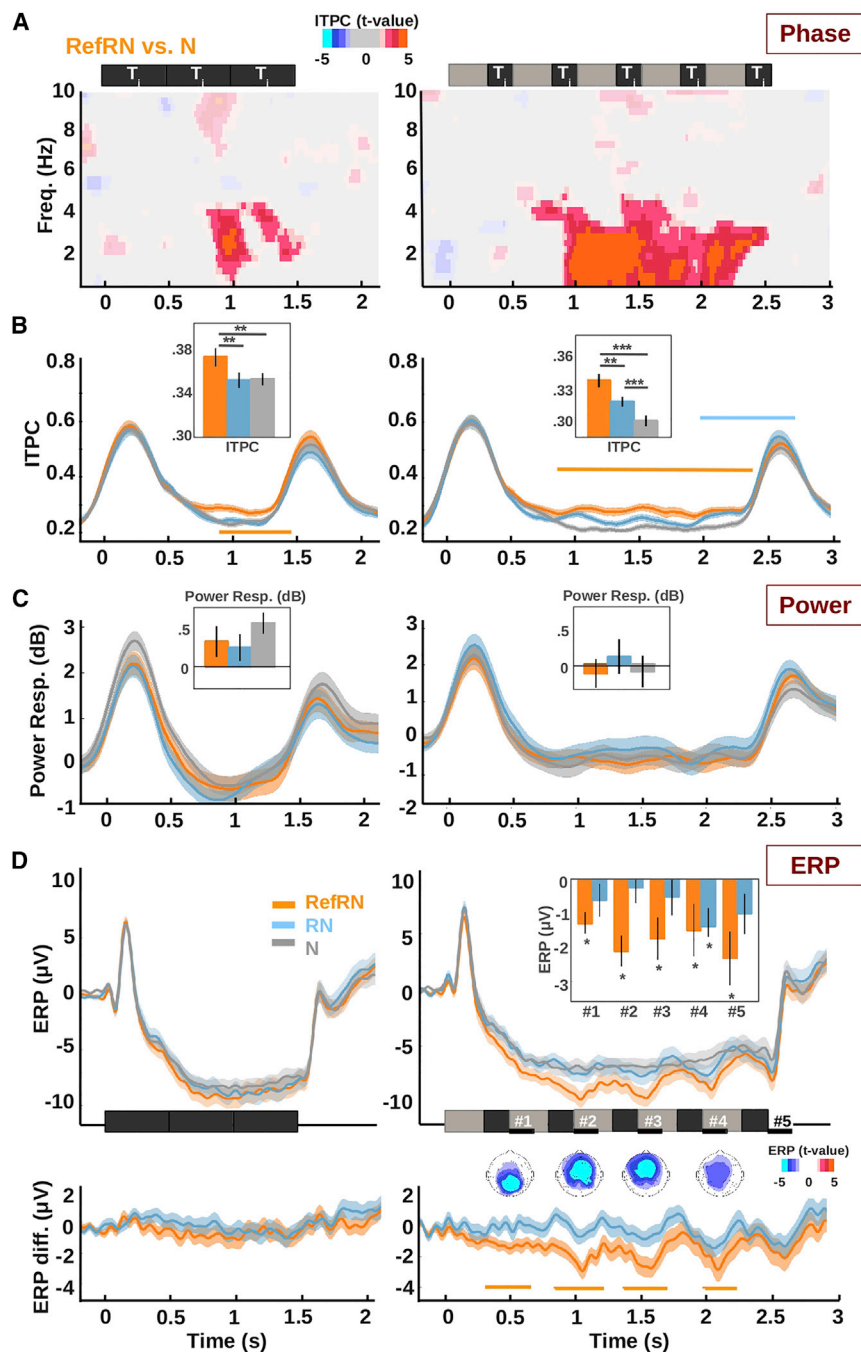
Figure 2. Electrophysiological Markers

Diffuse and compact conditions (experiment 1) are presented on the left and right columns, respectively.

(A) Time-frequency distribution of the increase of inter-trial phase coherency (ITPC) for RefRN compared to N trials (t values from uncorrected paired t tests across 42 blocks). The transparency mask shows clusters surviving a Monte-Carlo permutation test (Monte-Carlo p value < 0.05). Here and below results are averages for the ten most-responsive auditory electrodes (Figure S3A).

(B) Average ITPC in the 0.5–5 Hz region of interest for RefRN (orange), RN (blue), and N (gray). Horizontal colored lines show significant clusters for diffuse (RefRN versus N: [800, 1,400] ms; Monte-Carlo p value < 0.005) and compact (RefRN versus N: [800, 2,400] ms, Monte-Carlo p value < 0.0001; RN versus N: [2,000, 2,700] ms, Monte-Carlo p value < 0.0001) conditions. Insets show the mean ITPC further averaged over stimulus duration. Stars indicate the significance level of paired comparisons between conditions (paired t tests; ns: p > 0.05; *p < 0.05; **p < 0.01; ***p < 0.005).

(C) Power response in the 0.5–5 Hz region of interest, averaged across blocks. Insets show the mean power further averaged over stimulus duration. No significant difference could be observed between trial types.

(D) Evoked related potentials (ERPs) (top) and difference waves (RefRN or RN minus N; bottom). No statistical difference was observed between trial types for the diffuse condition. For the compact condition, averaging ERPs amplitude after repeated snippets (inset) revealed larger negativities for RefRN and RN compared to N (paired t tests; *p < 0.05). Difference waves also showed significant clusters (Monte-Carlo p value < 0.05, with topographies of t values also plotted). Note that the first cluster for the RefRN versus N comparisons start right after the first target onset. Error bars on insets and shaded areas around curves indicate SEM computed across blocks.

learning: ITPC (Figures 2A and 2B); EEG power (Figure 2C); and ERPs (Figure 2D).

For the diffuse condition (Figure 2, left; see legend for statistical tests), we observed higher ITPC for RefRNs in a [0.5, 5] Hz range. When averaging ITPC in this frequency range, only RefRN showed an increase compared to the N baseline. Further averaging ITPC over the whole stimulus duration (Figure 2B, inset) confirmed that the effect was restricted to RefRN. Applying the same analyses to power responses did not reveal any significant difference across conditions. Finally, we estimated ERPs time locked to stimuli onsets. We did not observe any difference

across stimulus types. So far, results for ITPC, power, and ERPs fully replicate the MEG findings of [5].

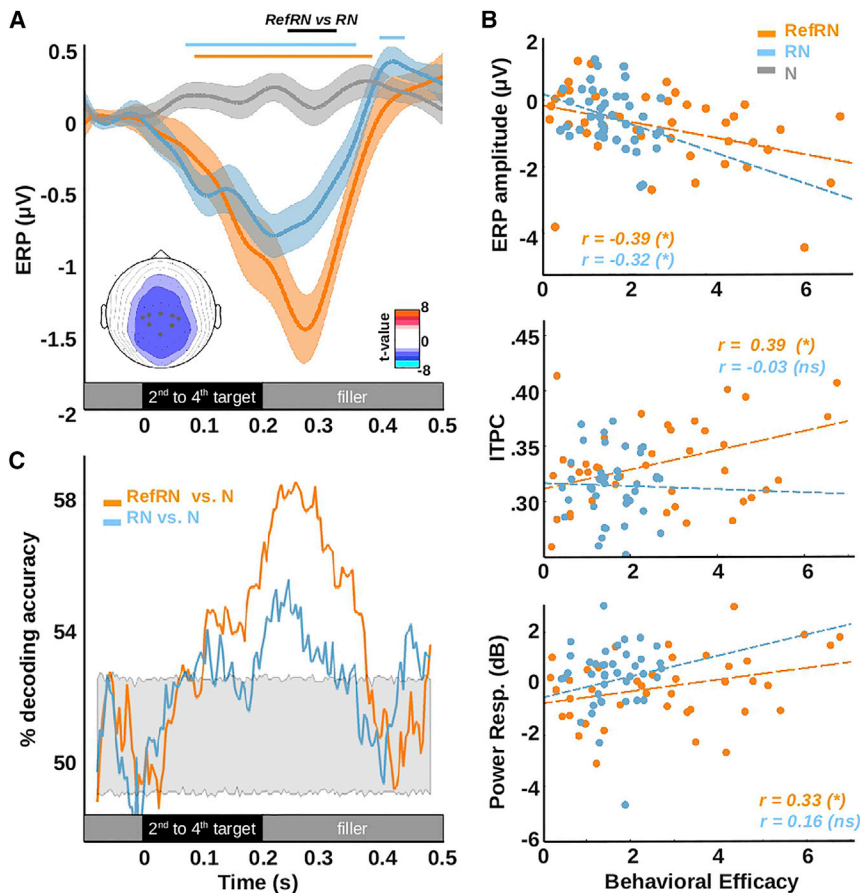For the compact condition (Figure 2, right), the same analyses were performed. Again, there was an increase in ITPC for RefRN compared to N. Averaging ITPC over the low-frequency range revealed a significant cluster for RefRN compared to N and here also for RN compared to N. As the noise snippets for RN were different from one trial to the next, this shows that across-trial phase patterns cannot be specific to a noise snippet. The power analysis did not reveal any difference across stimulus types. Finally, and crucially, there were clear modulations of the ERPs. Averaging ERPs amplitude after each repetition revealed consistent negative potentials for RefRN and RN (Figure 2D, inset). Remarkably, within the RefRN trials, ERPs were observed for each presentation of the repeated snippet, including the very first one (before any within-trial repetition).

Figure 3. Correlation of Neural Markers to Behavioral Performance for the Compact Condition

(A) ERPs were averaged across repetition epochs ([−100, 500] ms window from target onset; second to fourth within-trial target occurrences). Horizontal lines indicate significant clusters when comparing RefRN with N (orange; [128, 364] ms), RN with N (blue; negative: [72, 328] ms; positive: [388, 472] ms), and RefRN with RN (black; [244, 316] ms) trials (Monte-Carlo p values < 0.05). Shaded areas indicate SEM across blocks. The inset shows the topographical map of the differences between RefRN and N expressed as t values (paired t tests on averaged ERP amplitude extracted over a [100, 400]-ms window; n = 42 blocks). Non-significant t values (p > 0.05/65; Bonferroni correction) were set to white.

(B) Correlation of behavioral efficacy with ERP amplitude (top), phase coherence (ITPC; middle), and EEG power (bottom) for RefRN (orange) and RN (blue) trials. Pearson's correlation coefficients (r) were computed for RefRN and RN conditions separately and displayed on scatterplots along their statistical significance level (ns: p > 0.05; *p < 0.05; **p < 0.01; ***p < 0.005). Orange and blue dashed lines show the linear fit estimated for RefRN and RN conditions, respectively.

(C) Experimental conditions (RefRN versus N [orange]; RN versus N [blue]; Supplemental Experimental Procedures) were decoded, at the single-trial level, using a logistic regression on the ERPs displayed in (A). Gray area denotes the chance level obtained through permutations of trial types (n = 1,000). Decoding values above this gray area are higher than 95% of random values.

Therefore, such ERPs cannot be markers of within-trial repetition only. We also observed significant ERPs for RN trials but only after several within-trial presentations of the repeated snippet.

In summary, ERPs were observed in response to a noise snippet if and only if the same snippet had been heard before, within (RN) or across (RefRN) trials. The appearance of ERPs was extremely rapid, as they developed within five presentations of a novel noise snippet in RN trials. Such time-locked ERPs occurred without any discontinuities in sounds' amplitude or any other short-term statistics. To illustrate this point, we ran the stimuli through a peripheral auditory model (Figure 1B). The simulation showed that there were no obvious landmarks in RN/RefRN, at least not of the kind known to produce auditory ERPs before learning [3]. To stress that the "events" producing the ERPs were related to past experience, we term such responses "memory-evoked potentials" (MEPs).

MEPs Are Sensory Correlates of Behavioral Performance

To further characterize MEPs, we averaged responses time locked to the RN snippets for the compact condition (Figure 3A). Both RefRN and RN induced clear MEPs with a latency of about 100 ms. The MEPs' topography was very similar to the N1 topography (Figure S3A), but their broad, mostly negative waveform differed from a standard N1-P2 complex [3]. However, such topography and waveform are consistent with a superposition

of time-jittered N1-P2 complexes (see model below). The MEPs were larger for RefRN compared to RN. Nonetheless, after amplitude normalization, the waveforms and topographies became identical (Figure S3C). This suggests a common origin: for RN, MEPs could indicate the emergence of a mnesic trace toward the end of the trial, whereas for RefRN, the same mnesic trace would be re-activated from the very first snippet presentation and then reinforced by subsequent presentations.

If this unified account were correct, MEPs should always correlate with behavioral performance. This is exactly what was found: amplitude correlated with BE for both RefRN and RN (Figure 3B). We further tested whether MEPs could differentiate between stimuli on a trial-by-trial basis, using a logistic regression on the MEPs amplitude (Figure 3C). We obtained significant decoding as early as 100 ms post-presentation, supporting a sensory interpretation. The decoding accuracy was modest, but note that all sounds in this analysis were statistically exactly the same: a single epoch of white noise. Still, MEPs carried information about past experience, on a trial-by-trial basis.

Learning Noise without Paying Attention

So far, listeners were instructed to detect noise repetitions, so at least part of the MEPs could have been caused by attentional modulation. We tested this hypothesis in a supplemental experiment (Figure S2). Listeners were not asked to detect repetitions but rather, had to perform a distracting auditory task (detection
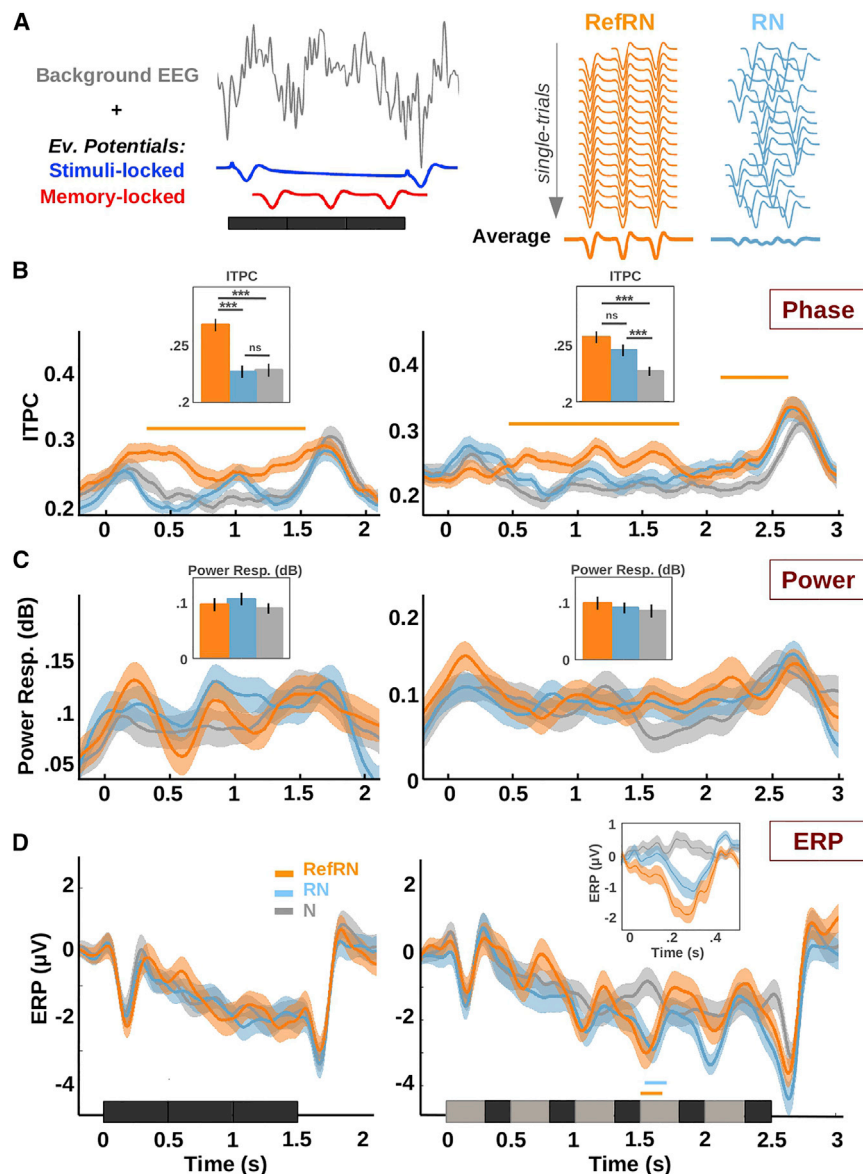
Figure 4. Model Simulations

(A) Illustration of the model's architecture. Background EEG was first synthesized (gray curve; Supplemental Experimental Procedures), and evoked potentials were added at the onsets and offsets of acoustic energy (stimulus-locked; dark blue). If a noise snippet had been heard before, an additional ERP was added (memory-locked; red), with a random onset time. The onset time was then fixed for subsequent presentations of the same noise snippet. The illustration for the diffuse condition (right) shows that, by construction, RefRNs were associated with perfectly synchronous potentials throughout a block, as they contained the same noise snippets across trials, whereas an RN trial contained only two potentials (after the first repetition epoch of each trial) with time jitter across trials. Forty-two blocks were simulated with a signal-to-noise ratio matching the empirical data set (see Supplemental Experimental Procedures and Figure S4).

(B–D) Analyses of the simulated data; format as Figure 2. Inset of (D) shows the target-locked ERPs as in Figure 3A. Colored lines denote significant clusters for the RefRN versus RN (orange) and RN versus N (blue) comparisons (Monte-Carlo p values < 0.05; n = 42 simulated blocks). In the insets, stars indicate the significance level (paired t tests; ns: $p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.005$).

gered and its onset time would remain exactly the same. But a different noise or listener would result in an ERP with a different, random onset time. As a result, for RN, the evoked activity was shifted across trials and shifted across blocks, whereas for RefRN, the evoked activity was fixed across trials and shifted across blocks.

We analyzed the simulated data (Supplemental Experimental Procedures) in the same way as the EEG recordings. The model replicated all of the main findings (Figures 4B–4D). In particular, no ERPs were observed in the diffuse condition, as the memory-locked N1-P2 averaged out across blocks, due to the 500-ms onset-time jitter. Time-locked ERPs similar to MEPs were observed only for the compact condition, as the 200-ms onset-time jitter was too short for N1-P2 components to fully overlap and cancel out. The peculiar shape of the MEPs itself was reproduced by the additive model (Figure 4D, inset).

DISCUSSION

We used acoustic noise to probe the neural bases of auditory perceptual learning. Our results outline a simple mechanistic account of how initially nondescript, random sounds may acquire perceptual uniqueness. Through exposure, with or without focused attention, rapid plasticity creates sensory selectivity to subtle acoustic details within a specific noise pattern. Such

of amplitude modulations) [5]. In addition, RN or RefRN sequences were embedded in 8 min of continuous running noise: there was no amplitude-onset cue to signal that a new "learnable" sequence had begun, thus removing endogenous and exogenous attentional cues. Still, clear MEPs were observed, remarkably similar to those of the main experiment (Figure S3B).

A Simple Model of Memory-Evoked Responses

A possible interpretation for the MEPs is that they were triggered by acoustic events within the noise, which only became perceptually salient after learning. We implemented this idea in a simple quantitative model (Figure 4A). Whenever a snippet of noise had been heard before, we injected an ERP in the EEG waveform, with a canonical shape (N1-P2) [3] and random onset time for each "listener" and noise. This was intended as an idealized version of one-shot learning: whenever the same noise would be heard again by the same listener, an ERP was invariably trig-

details are localized in time, idiosyncratic, and only become salient after perceptual learning.

This account clarifies the puzzling issue of what is learned within a noise. RN is introspectively reported as containing short "rasping, clanking" perceptual events [23, 28]. Behavioral data already suggested that those events differed across listeners for the same noise [2, 22, 23] and thus could not be unambiguously traced back to acoustic landmarks. Our EEG data support this idea. Noise does contain short-term variations, which could be reflected by cortical activations [29, 30]. However, if evoked potentials were due to a passive transmission of acoustic landmarks, all repeated snippets should have been equally signaled. Instead, we found that evoked potentials developed over time, correlated closely with behavior, and were consistent with a model of idiosyncratic perceptual learning.

These MEPs were interpreted as the superposition of standard N1-P2 complexes. The N1-P2 complex has been associated to perceptual changes within constant-amplitude stimuli [31], and it can be modulated by repeated exposure [32, 33]. Here, we demonstrated not only a modulation of N1-P2 on a much-faster timescale but also the appearance of such an ERP where there was none before learning. As the planum temporale is one of the cortical sources of the N1-P2 complex [34], our results also advocates for a role of this secondary auditory structure in rapid plasticity and sensory memory. This is consistent with fMRI results using a noise-learning paradigm [6] or, more classically, with mismatch-negativity studies [26, 35].

Computationally, the learning of discriminant patterns within noise could be achieved through established plasticity mechanisms such as spike-time-dependent plasticity (STDP). In STDP models, repeated exposure to random afferent patterns almost inevitably leads to pattern-specific selectivity at the single-neuron [17] or small-network level [18]. Experimentally, however, we recorded ERPs on scalp electrodes, which must involve relatively broad neural networks. So, was the code local or global? A possibility is that the scalp potentials were the outcome of a cascade of neural events, initially triggered by a sparse mnesic trace [36] and then amplified by perceptual awareness [37]. Indeed, perceptual awareness of a target tone embedded within a stochastic masker is associated with an N1-like ERP [21]. So, even if our experimental measure was at the network level, we argue that, altogether, the data and model suggest a highly local neural code for experience-dependent changes induced by the perceptual learning of noise.

Functionally, sharp neural selectivity to past sensory experiences would help the auditory system distinguish previously heard sounds from truly novel ones. More generally, it could aid learning about frequently encountered natural sounds. In this respect, the search for generic timbre dimensions useful for the identification of sound sources has proven surprisingly elusive [2, 38, 39]. The present results suggest that this may be because source identification is shaped as much by idiosyncratic experience as by acoustic properties.

EXPERIMENTAL PROCEDURES

A brief description of experimental procedures can be found in the Results section. A complete description can be found in the Supplemental Experimental Procedures. The experimental protocols were approved by the local ethical committee (Conseil d'Evaluation Ethique pour les Recherches en Santé, University Paris Descartes).

REFERENCES

1. Gilbert, C.D., Sigman, M., and Crist, R.E. (2001). The neural basis of perceptual learning. Neuron 31, 681–697.

2. Agus, T.R., Thorpe, S.J., and Pressnitzer, D. (2010). Rapid formation of robust auditory memories: insights from noise. Neuron 66, 610–618.

3. Picton, T.W. (2010). Late auditory evoked potentials: changing the things which are. In Human Auditory Evoked Potentials (Plural Publishing), pp. 335–398.

4. Guttman, N., and Julesz, B. (1963). Lower limits of auditory periodicity analysis. J. Acoust. Soc. Am. 35, 610.

5. Luo, H., Tian, X., Song, K., Zhou, K., and Poeppel, D. (2013). Neural response phase tracks how listeners learn new acoustic representations. Curr. Biol. 23, 968–974.

6. Kumar, S., Bonnici, H.M., Teki, S., Agus, T.R., Pressnitzer, D., Maguire, E.A., and Griffiths, T.D. (2014). Representations of specific acoustic patterns in the auditory cortex and hippocampus. Proc. Biol. Sci. 281, 20141000.

7. Mizrahi, A., Shalev, A., and Nelken, I. (2014). Single neuron and population coding of natural sounds in auditory cortex. Curr. Opin. Neurobiol. 24, 103–110.

8. Agus, T.R., and Pressnitzer, D. (2013). The detection of repetitions in noise before and after perceptual learning. J. Acoust. Soc. Am. 134, 464–473.

9. O'Toole, A.J., and Kersten, D.J. (1992). Learning to see random-dot stereograms. Perception 21, 227–243.

10. Gold, J.M., Aizenman, A., Bond, S.M., and Sekuler, R. (2014). Memory and incidental learning for visual frozen noise sequences. Vision Res. 99, 19–36.

11. Thorpe, S.J. (2011). Grandmother cells and distributed representations. In Visual Population Codes: Toward a Common Multivariate Framework for Cell Recording and Functional Imaging, G. Kreiman, and N. Kriegeskorte, eds. (MIT Press), pp. 23–51.

12. Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat. Neurosci. 6, 1216–1223.

13. Schreiner, C.E., and Polley, D.B. (2014). Auditory map plasticity: diversity in causes and consequences. Curr. Opin. Neurobiol. 24, 143–156.

14. Elliott, T.M., Hamilton, L.S., and Theunissen, F.E. (2013). Acoustic structure of the five perceptual dimensions of timbre in orchestral instrument tones. J. Acoust. Soc. Am. 133, 389–404.

15. McDermott, J.H., Schemitsch, M., and Simoncelli, E.P. (2013). Summary statistics in auditory perception. Nat. Neurosci. 16, 493–498.

16. Bathellier, B., Ushakova, L., and Rumpel, S. (2012). Discrete neocortical dynamics predict behavioral categorization of sounds. Neuron 76, 435–449.

17. Masquelier, T., Guyonneau, R., and Thorpe, S.J. (2008). Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains. PLoS ONE 3, e1377.

18. Klampfl, S., and Maass, W. (2013). Emergence of dynamic memory traces in cortical microcircuit models through STDP. J. Neurosci. 33, 11515–11529.

19. Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. 15, 511–517.

20. Luo, H., and Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. Neuron 54, 1001–1010.

21. Gutschalk, A., Micheyl, C., and Oxenham, A.J. (2008). Neural correlates of auditory perceptual awareness under informational masking. PLoS Biol. 6, e138.

22. Kaernbach, C. (1992). On the consistency of tapping to repeated noise. J. Acoust. Soc. Am. 92, 788–793.

23. Kaernbach, C. (1993). Temporal and spectral basis of the features perceived in repeated noise. J. Acoust. Soc. Am. 94, 91–97.

24. Moore, B.C.J. (2003). Temporal integration and context effects in hearing. J. Phonetics 31, 563–574.

25. Kaernbach, C. (2004). The memory of noise. Exp. Psychol. 51, 240–248.

26. Berti, S., Schröger, E., and Mecklinger, A. (2000). Attentive and pre-attentive periodicity analysis in auditory memory: an event-related brain potential study. Neuroreport 11, 1883–1887.

27. Macmillan, N.A. (2005). Detection Theory: A User's Guide, Second Edition (Mahwah, N.J.: Lawrence Erlbaum Associates).

28. Julesz, B., and Guttman, N. (1963). Auditory memory. J. Acoust. Soc. Am. 35, 1895.

29. Barker, D., Plack, C.J., and Hall, D.A. (2012). Reexamining the evidence for a pitch-sensitive region: a human fMRI study using iterated ripple noise. Cereb. Cortex 22, 745–753.

30. Steinmann, I., and Gutschalk, A. (2012). Sustained BOLD and theta activity in auditory cortex are related to slow stimulus fluctuations rather than to pitch. J. Neurophysiol. 107, 3458–3467.

31. Krumbholz, K., Patterson, R.D., Seither-Preisler, A., Lammertmann, C., and Lütkenhöner, B. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cereb. Cortex 13, 765–772.

32. Ross, B., Jamali, S., and Tremblay, K.L. (2013). Plasticity in neuromagnetic cortical responses suggests enhanced auditory object representation. BMC Neurosci. 14, 151.

33. Tremblay, K.L., Ross, B., Inoue, K., McClannahan, K., and Collet, G. (2014). Is the auditory evoked P2 response a biomarker of learning? Front. Syst. Neurosci. 8, 28.

34. Näätänen, R., and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24, 375–425.

35. Näätänen, R., Jacobsen, T., and Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. Psychophysiology 42, 25–32.

36. Hromádka, T., Deweese, M.R., and Zador, A.M. (2008). Sparse representation of sounds in the unanesthetized auditory cortex. PLoS Biol. 6, e16.

37. Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., and Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends Cogn. Sci. 10, 204–211.

38. Patil, K., Pressnitzer, D., Shamma, S., and Elhilali, M. (2012). Music in our ears: the biological bases of musical timbre perception. PLoS Comput. Biol. 8, e1002759.

39. Leaver, A.M., and Rauschecker, J.P. (2010). Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. J. Neurosci. 30, 7604–7612.